

Weekly Report

2013-10-27

Feiran Wu

1 Introduction

Our job in this week can be summarized into two parts. One is for Information entropy in the use of air pollution data. The other is for the paper organization and reading.

1 Information entropy on the extraction of time-series event

The pinus view in “A Visual Analytics Approach to Multi-scale Exploration of Environmental Time Series” is a method for multi-scale single time sequence visualization. We implement it on air pollution data as Figure 1 shown. Comparing to the Figure 2, the result is similar, the pinus view like a combination of many triangles. With this method, we can select on pinus view and extract the corresponding sub-time-series.

The problem is that, now we only select all the attributes then calculate the information entropy. If we divide all the attributes into n groups, there are n pinus views. Which means we need to design a multi-pinus-view schema to make event selection and comparison. Now Fangzhou is programming for a simple select and event highlight version.

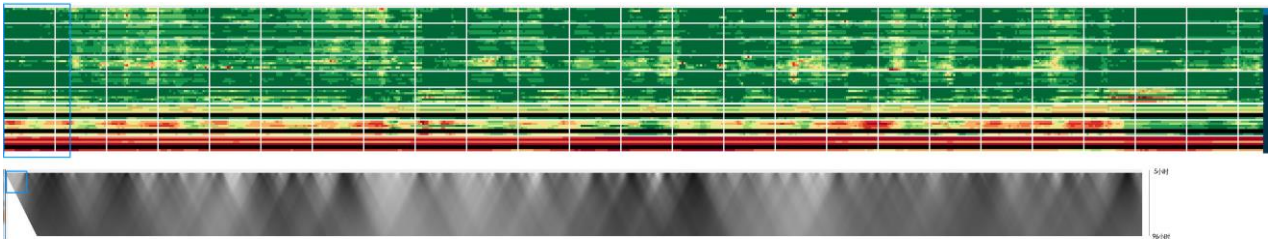


Figure 1 the bitmap view and the pinus view for air pollution data.

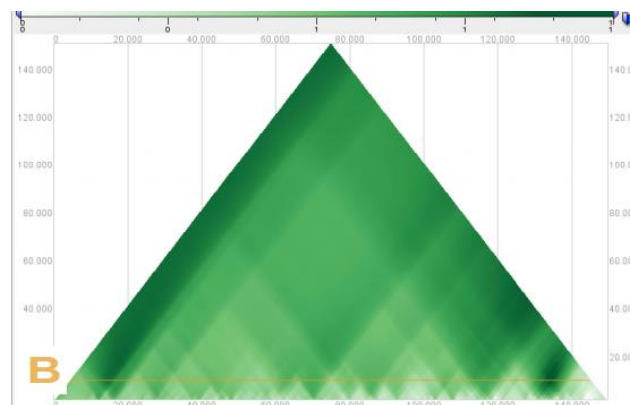


Figure 2 the visualization result in original paper for multi-scale information entropy.

2 Phase summary of paper reading

I categorize the candidate event extraction and co-occurrence method into two groups. One is the statistical approach which removes the time and consider the time points selected as random sampling (e.g. Information entropy, mutual information, CCA). The other is the shape-matching approach (e.g. time series motif discovery method (random projections, viztree)).

The concept “co-occurrence”, I apt to consider it as a multi-dimension pattern query problem. Now I have two guidances to find the “co-occurrence” events. One is “Partition”. Firstly we find a large multi-dimensional event. Then divide this event into several parts. Every part corresponding to several dimensions. The other is “Combination”. Grouping the similar dimensions according to priori knowledge or automatic algorithm and extracting the event. Then construct the relationship among these events.

The goal of our research is still the time-series event extraction and event relationship analysis. Next week I will continue to think about this question, while I stay in Guilin for 4 days to to deliver the academic presentation.